Structural and functional studies of the biotin protein ligase from Aquifex aeolicus reveal a critical role for a conserved residue in target specificity

Cite as:
Tron, C. M., McNae, I. W., Nutley, M., Clarke, D. J., Cooper, A., Walkinshaw, M. D., Baxter, R. L., & Campopiano, D. J. (2009). Structural and functional studies of the biotin protein ligase from Aquifex aeolicus reveal a critical role for a conserved residue in target specificity. *Journal of Molecular Biology*, 387(1), 129-46.

Manuscript received: 15/09/2008; Accepted: 29/12/2008; Article published: 22/01/2009

Structural and Functional Studies of the Biotin Protein Ligase from *Aquifex aeolicus* Reveal a Critical Role for a Conserved Residue in Target Specificity**

Cecile M. Tron,[1,†] Iain W. McNae,[2,†] Margaret Nutley,[3] David J. Clarke,[1] Alan Cooper,[3] Malcolm D. Walkinshaw,[2] Robert L. Baxter[1] and Dominic J. Campopiano[1]

[1]EaStCHEM, School of Chemistry, Joseph Black Building, University of Edinburgh, West Mains Road, Edinburgh, EH9 3JJ, UK.

[2]Structural Biochemistry, The University of Edinburgh, Michael Swann Building, King's Buildings, Edinburgh, EH9 3JR, Scotland, UK.

[3]WestChem Department of Chemistry, University of Glasgow, Glasgow G12 8QQ, Scotland, UK.

[*]Corresponding authors; D.J.C. e-mail: Dominic.Campopiano@ed.ac.uk, tel: 44 131 650 4712, fax: 44 131 650 4743; and R.L.B. e-mail: Robert.Baxter@ed.ac.uk; tel: 44 131 650 4708, fax: 44 131 650 4743.

[**]We are grateful to the UK BBSRC (Biotechnology and Biological Sciences Research Council) for their support of C. M. T. and the ITC facility in Glasgow. C. M. T. and D. J. C. also thank the University of Edinburgh for post-graduate funding. Thanks also to Chris Brown for his advice on the early stages of this project.

[†]C.M.T. and I.W.M. contributed equally to this work.

Supporting information:
Supplementary data for this article can be found online at http://dx.doi.org/10.1016/j.jmb.2008.12.086

Keywords:
biotin protein ligase; biotinylation; post-translation modification; X-ray crystallography; Isothermal titration calorimetry

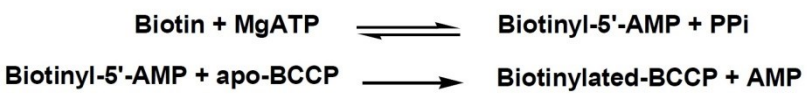
Abbreviations:
BPL, biotin protein ligase; BCCP, biotin carboxyl carrier protein; BCCPΔ67, C-terminal biotinyl domain; HCS, holocarboxylase synthetase; birA; *E. coli* biotin protein ligase; ITC, Isothermal titration calorimetry.

Abstract

Biotin protein ligase (BPL, EC: 6.3.4.15) catalyses the formation of biotinyl-5'-AMP from biotin and ATP, and the succeeding biotinylation of the biotin carboxyl carrier protein (BCCP). We describe the crystal structures at 2.4 Å resolution of the class I BPL from the hyperthermophilic bacteria *Aquifex aeolicus* (*Aa*BPL) in its ligand-free form and in complex with biotin and ATP. The solvent-exposed β- and γ-phosphates of ATP are located in the inter-subunit cavity formed by the N- and C-terminal domains. The Arg40 residue from the conserved GXGRXG motif is shown to interact with the carboxyl group of biotin and to stabilise the α- and β- phosphates of the nucleotide. The structures of the mutant *Aa*BPL R40G in both the ligand-free and biotin-bound forms reveal that the mutated loop has collapsed, thus hindering ATP binding. Isothermal titration calorimetry experiments indicate that the presence of biotin is not required for ATP binding to wildtype *Aa*BPL in the absence of Mg^{2+} ions and the binding of biotin and ATP has been determined to occur *via* a random but cooperative process. The affinity for biotin is relatively unaffected by the R40G mutation. In contrast, the thermodynamic data indicate that binding of ATP to *Aa*BPL R40G is very weak in the absence or presence of biotin. The *Aa*BPL R40G mutant remains catalytically active but shows poor substrate specificity - mass spectrometry and western blot studies reveal that the mutant biotinylates the target *A. aeolicus* BCCPΔ67 fragment but also BSA, and is subject to self-biotinylation.

Introduction

The vitamin biotin plays key metabolic roles when covalently bound to essential carboxylases and decarboxylases involved in lipogenesis, glucogenesis, and amino acid degradation.[1] The enzymatic attachment of the biotin co-factor is catalysed by biotin protein ligase (BPL, EC: 6.3.4.15), also known as holocarboxylase synthetase (HCS). BPL is exceptionally specific, recognising less than five biotin-dependant enzymes in most eukaryotes and in bacteria only the multi-subunit protein acetyl-CoA carboxylase (ACC) is biotinylated.[2] In the functional ACC complex, the biotin prosthetic group is linked *via* an amide linkage to a conserved lysine residue of the biotin carboxyl carrier protein (BCCP) domain.[3] Biotin mediates (through its N_1 carboxylated form) the transfer of CO_2 from hydrogen carbonate to acetyl-CoA to afford malonyl-CoA.[4] The biotinylation reaction catalysed by BPL is Mg^{2+} and ATP dependant and occurs in two steps (Scheme 1). The first step generates the reactive intermediate biotinyl-5'-AMP and in the second step, the N_ε of the target lysine of BCCP reacts with the activated carboxyl group to form biotinylated-BCCP with release of AMP.[5]



Scheme 1. The biotinylation reaction catalysed by BPL.

Mechanistically, BPL belongs to the family of adenylating enzymes which includes lipoate protein ligase (LplA) and aminoacyl-tRNA synthetase (aaRS). Analysis of the crystal structures of *E. coli* BPL, BirA, and LplA shows that the two enzymes are close structural homologs.[6-8] Structural similarities have also been previously reported between BirA and class II aminoacyl-tRNA synthetases and these suggest that despite poor sequence identity, the three enzymes could have arisen from a common ancestor.[7,8] The BPL enzymes have been further classified into four sub-classes, according to their domain structures and their ability to control biotin metabolism.[9] *E. coli* BirA is a bifunctional class II BPL which has regulatory properties as well as biotin-transfer activity. BirA normally acts as a ligase after converting biotin into biotinyl-5'-AMP, but in the absence of substrate apo-BCCP, the BirA:biotinyl-5'-AMP complex dimerizes and represses biotin biosynthesis by binding of the N-terminal domains to the bidirectional O/P region of the biotin operon (which encodes *bio*ABCFD).[10,11] The biotin sensing/birA regulatory circuit has been used as a paradigm for studies of the control of metabolism by a repressor protein controlled by a small molecule co-regulator.[12]

The structure of BirA has been reported in its ligand-free form, in complex with biotin and with the biotinyl-5'-AMP analogue, biotinol-5'-AMP.[13-16] These structures show three distinct domains: a N-terminal DNA-binding domain, a large central catalytic domain, and a small C-terminal domain for which the exact function remains unclear. The structure of monomeric, ligand-free BirA reveals four disordered surface loops in the catalytic domain and shows that disorder-to-order transitions occurring upon biotin binding induce dimerization. In the structure of BirA in complex with biotin, the biotin binding loop protects the biotin moiety from the solvent while two more loops located at the dimer interface are also ordered. In the structure of the BirA:biotinol-5'-AMP complex, the last peptide segment forms the adenylate-binding loop and completes the dimer packing. In BirA, the binding of the substrates and the formation of biotinyl-5'- AMP occur *via* an ordered sequential process with ATP binding after biotin and subsequent dimerisation.[17]

The crystal structures of *Pyroccocus horikoshii* OT3 BPL (*Ph*BPL) have been determined in the ligand-free state and in complex with biotin, ADP, and biotinyl-5'- AMP.[18] The class I *Ph*BPL lacks the N-terminal DNA-binding domain and acts solely as a ligase but the overall fold of *Ph*BPL has high structural similarity with the catalytic and C-terminal subunits of BirA. The structure of ligand-free *Ph*BPL indicates that only the biotin binding loop is disordered and that after substrate binding, the intermediate biotinyl-5'-AMP is stabilised within the active site *via* hydrogen-bonds with the loop. The formation of a BPL:BCCP complex appears to be required prior to the transfer of the biotin moiety onto the target lysine to ensure the specificity of the enzymatic biotinylation reaction. Most recently, the structures of the single and double mutants, *Ph*BPL R48A and *Ph*BPL R48A K111A, as well as their complexes with the corresponding C-terminal BCCP domain (PhBCCPΔN76) revealed the residues responsible for complex formation and biotin transfer in this hypothermophilic species.[19] A model of the *E. coli* complex has also been assembled using the structures of BirA and the *E. coli* BCCP87 biotinyl domain, and the complex formed between *Aquifex aeolicus* BPL and the C-terminal fragment BCCPΔ67 has been characterised by chemical cross-linking.[20,21]

The conserved glycine-rich motif, GXGRXG, present in all BPLs (e.g. $_{45}$GHGRLN$_{50}$ in *P. horikoshii*), is located on the turn of the biotin binding loop in the structures of *E. coli* BirA and *Ph*BPL. In BirA this sequence motif, originally thought to be involved in binding the nucleotide portion of ATP, was shown to be involved in binding of biotin and the adenylate intermediate.[22] In contrast, the corresponding loop has been shown in *Ph*BPL to interact with biotin, ADP and biotinyl-5'-AMP.[18] The two mutants *Ph*BPL R48A and *Ph*BPL R48A K11A have been shown to catalyse the formation of holo-BCCPΔN76.[19] However, while the structure of the single *Ph*BPL mutant in complex with biotinyl-5'AMP has been determined showing near-identical interactions with the native enzyme, in the structure of *Ph*BPL R48A K11A, only biotin and the adenosine moiety are visible. In the BirA structures, the glycine-rich motif plays a role in binding both biotin and biotinol-5'-AMP.[14,16] Kinetic and thermodynamic studies carried out on single-site mutants of the $_{115}$GRGRRG$_{120}$ motif of BirA indicate that mutations within this motif alter ligand affinities and affect the dimerisation process.[22-24] Mutant BirA R118G has been shown to have lost its substrate specificity; *in vivo* and in *vitro* assays have shown that the mutant biotinylates itself as well as non-cognate acceptors.[25] While the mechanism involved in this *promiscuous* biotinylation is not understood, it has been suggested that it could arises as a result of dissociation of the BirA R118G:biotinyl-5'-AMP complex to give free biotin adenylate which reacts with available amino groups. Despite the crystallographic data obtained for the BirA and *Ph*BPL complexes, the role of the conserved arginine in the normal catalytic process remains unclear. The mechanistic importance of this residue is underscored by the observation that mutation of the corresponding arginine in the glycine motif of human holocarboxylase synthetase (R508W), which is known to alter its affinity for biotin, gives rise to multiple carboxylase deficiency syndrome.[26,27]

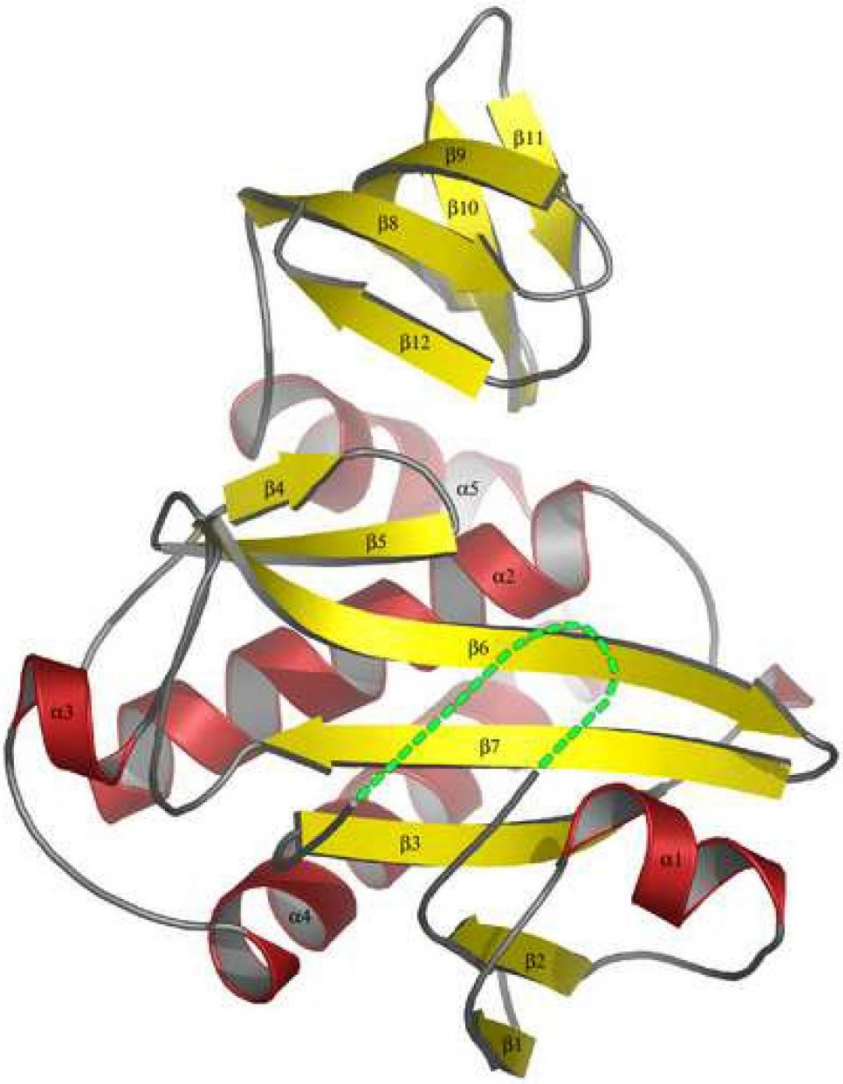
We have investigated the importance of the arginine 40 from the glycine-rich motif in the specificity of the biotinylation reaction catalysed by the *Aquifex aeolicus* BPL (*Aa*BPL). The X-ray structures of the class I *Aa*BPL in the ligand-free state and in complex with biotin and ATP are described and these allow us to propose a role for the Arg40 residue in substrate binding. Isothermal titration calorimetry experiments have been carried out with *Aa*BPL in order to investigate the substrate binding processes. Characterisation of the crystal structure of the mutant *Aa*BPL R40G in complex with biotin and complementary ITC experiments indicate that the conformation of the active site and ATP binding are strongly affected by the R40G mutation. Furthermore, streptavidin western blot studies and mass spectrometric analyses carried out with wildtype *Aa*BPL and mutant R40G have shown the importance of the conserved Arg40 in the catalytic process.

Results and Discussion

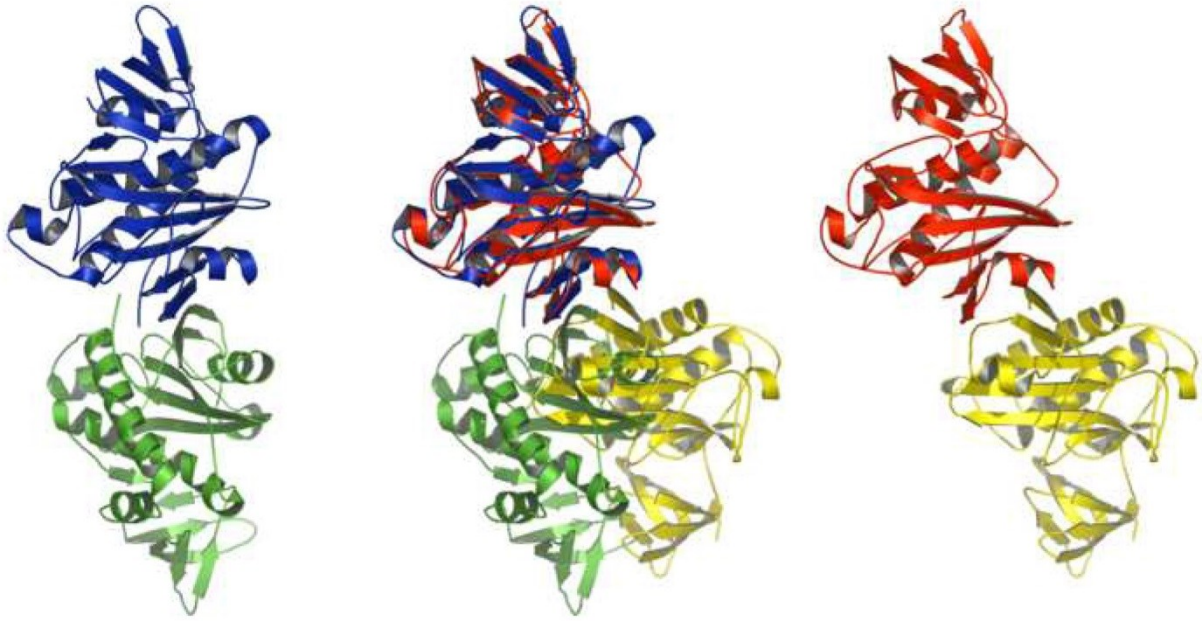
1. The Structure of A. aeolicus BPL

In the structure of the class I *Aa*BPL, the first 186 residues of the 26 kDa (233 aa) protein folds into a large N-terminal catalytic domain which contains seven β-strands and five α-helices (Figure 1a).[21] The smaller C-

terminal domain (47 aa) is formed by a mixed β-sheet of five strands. This two-domain topology is also observed in the structures of *E. coli* BirA and *P. horikoshii* BPL.[15,18] *Aa*BPL displays linear sequence similarities with BirA and *Ph*BPL of 20.9 % and 30.6 % sequence identity respectively. Analysis of the crystal structures of the three different BPLs reveals that similarities are also observed in their overall folds and secondary structures. Several conformational differences, mainly located on the loops involved in substrate binding, are observed between the active sites of the three BPLs. In the ligand-free structure of *Aa*BPL, residues 37 to 47 of the biotin binding loop (loop β_2-β_3) containing the $_{37}$GRGRLG$_{42}$ motif are disordered, and the active site formed by the strands β_3, β_6 and β_7 is solvent-exposed. The equivalent adenylatebinding loop of BirA, which only becomes organised upon ATP binding, is ordered in the ligand-free structure of *Aa*BPL and the ATP binding site is consequently more preformed.[14,16] Additionally, in the structure of *Aa*BPL, the surface loop connecting the β_6 and β_7 strands is shifted towards the binding site (Figure 1b).



(a)



(b)

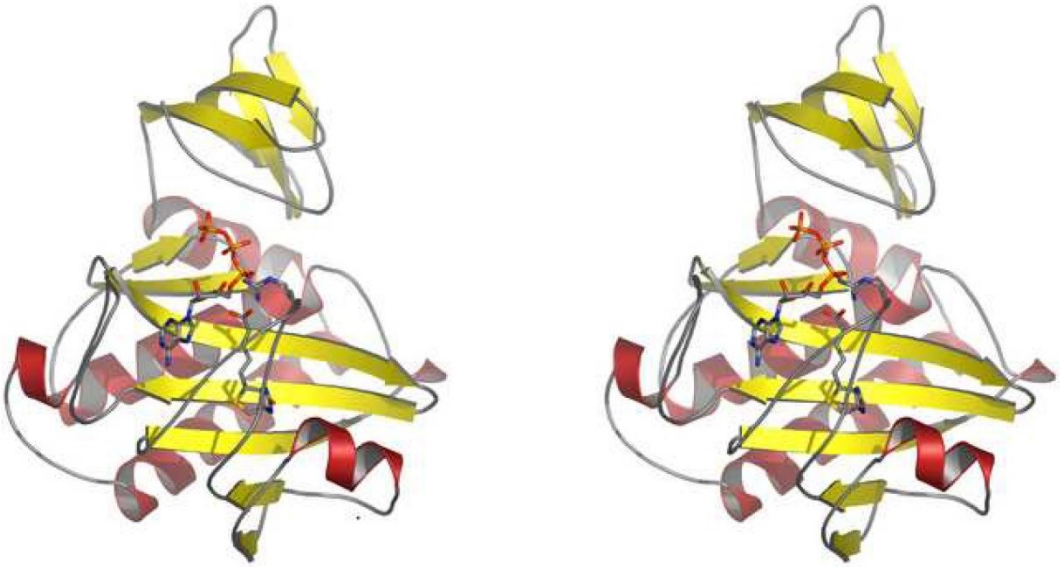
Figure 1. A cartoon ribbon representation of the apo form of *Aa*BPL and *Ph*BPL (PDB code 2DTO). (a) Monomer structure of apo form of *Aa*BPL showing secondary structure elements. The α-helices are coloured red and the β-strands are coloured yellow. Secondary structure elements are numbered according to their position in the protein sequence. The missing loop is represented by a broken green line. (b) Dimer comparison of *Ph*BPL and *Aa*BPL. Left-hand, respectively. Right-hand,$_6$ and β_7 strands is marked with an asterisk (*) and the colouring is as for individual images.

The class I *Ph*BPL is dimeric in solution and this dimer is also observed in the crystal form.[18] The β_1-strands of each *Ph*BPL monomer are hydrogen-bonded and form the dimer interface (Figure 1b). The crystals of *Aa*BPL belong to either the monoclinic space-group P21 or to the orthorhombic space-group $P2_12_12_1$ and, like the *P. horikoshii* enzyme, consist of two monomers per asymmetric unit related by a pseudo-two fold axis. However, in the structure of apo-*Aa*BPL the β_1-strands are located in close proximity but do not form an inter-subunit β-sheet (Figure 1b). The monomeric nature of *Aa*BPL observed in the crystal structures of the ligand-free and ligand-bound forms is also found in solution - supported by gel filtration experiments carried out in the absence and presence of substrates (data not shown).

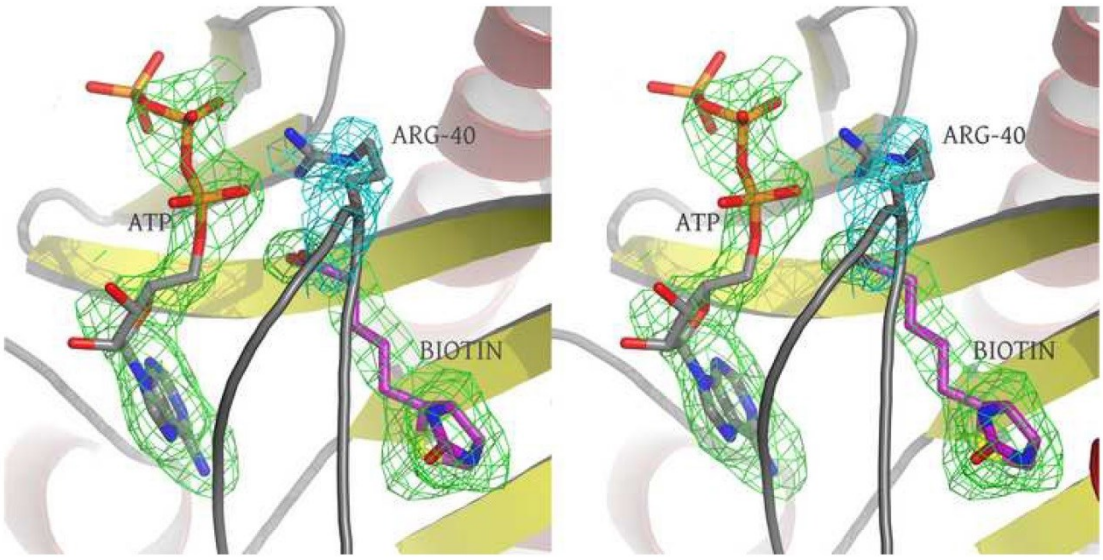
2. The biotin and ATP binding sites in AaBPL

The structure of *Aa*BPL complexed with biotin and ATP (*Aa*BPL:biotin:ATP) was determined at a resolution of 2.3 Å (Figure 2a). In the complex, residues 37-47, which form the biotin binding loop and act as a lid to the

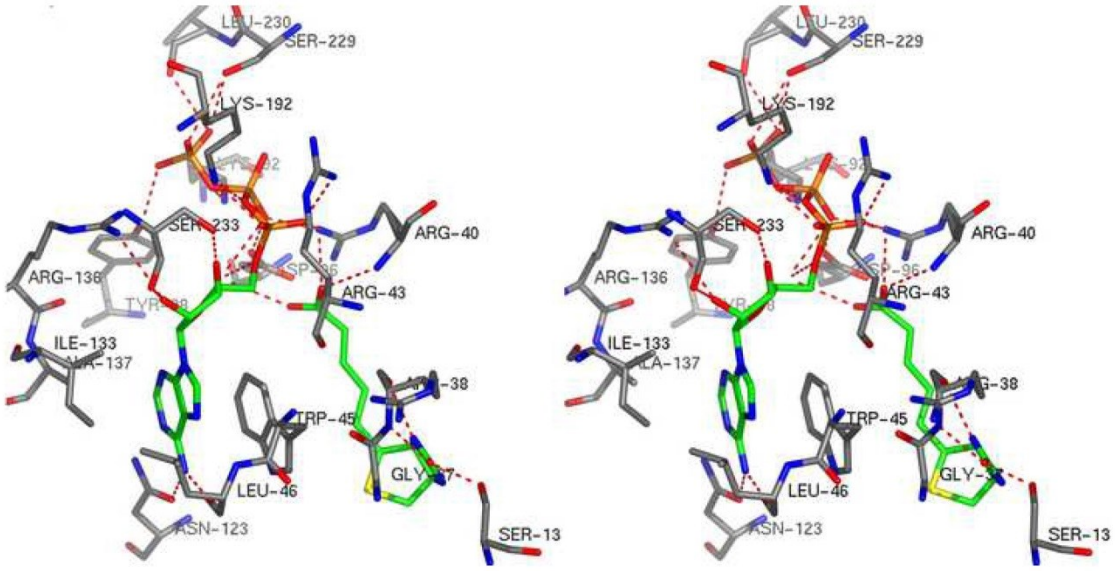
active site, and the two ligands are fully ordered in the electron density (Figure 2b) The loop stabilises the biotin and ATP moieties with strong hydrophobic and hydrophilic interactions. The tetrahydrothiophene ring of the biotin is accommodated in the interior of the binding pocket formed by residues Gly119 and Gly121 of the β_7-strand (Figure 2c) The nitrogen atoms of the biotin *ureido* group hydrogen-bond with the side chain of Gln34 and the backbone carbonyl of Arg38 from the biotin binding loop, as well as the Thr14 from the α_1-helix. In addition, the *ureido* carbonyl forms hydrogen-bonds with the hydroxyl group of Ser13 from the β_1- α_1 loop and the backbone NH of Arg38. The aliphatic tail of the biotin is held by a sandwich of two hydrophobic walls formed by the residues Gly106, Val107, and Leu108 from the β_6-strand on one side, and on the other side, the residues Gly37, Gly39, Trp45, and Leu46 from the biotin binding loop.



(a)



(b)



(c)

Figure 2. Stereo views of *Aa*BPL with bound biotin and ATP (*Aa*BPL:biotin:ATP). (a) A cartoon view of *Aa*BPL monomer, ATP, biotin and Arg40 are represented as sticks. (b) A close-up view of the active site with electron density around ATP, biotin and Arg40. The final omit $|F_o| - |F_c|$ is contoured at 2.5 σ . (Green chicken wire around ATP and BIOTIN, blue around Arg40) (c) Binding interactions of ATP and biotin within the *Aa*BPL protomer A active site. ATP and biotin are coloured with green carbons, H-bonds are shown as broken red lines. Residues (Lys192 and Ser233) from the other protomer are annotated with asterisks (*).

The biotin carboxylate is adjacent to the turn of the biotin binding loop and is strongly stabilised by hydrogen-bonds with the side chain of the Arg40 from the GRGRLG motif and have weak interactions with the Arg40 backbone NH and the side chain of Lys103 from the β_6-strand. The carboxylate group lies toward the outside of the active site, near the ATP binding pocket. We find that the oxygen atoms of the biotin carboxylate group and the ・-phosphate of ATP are located in close proximity (approximately 3.5 Å) (Figures 2b and c). Therefore, the nucleophilic attack of the carboxylate anion at the α- phosphate of ATP, resulting in the formation of the phosphoanhydride bond of biotinyl- 5'-AMP, induces only minor conformational changes to the active site. Comparison of the structures of the *Aa*BPL:biotin:ATP and *Ph*BPL:biotinyl-5'-AMP complexes shows that the orientations of the adenine ring and sugar moiety of ATP and the biotin within the active sites are similar in the two complexes.[18]

We find that polar and electrostatic interactions with residues from both the adenylatebinding loop (Asn123 and Ile133) and the biotin binding loop (Trp45 and Leu46) contribute to the correct positioning of the adenine ring of ATP in the active site (Figure 2c). The α-phosphate of ATP is stabilised by a network of hydrogen-bonding interactions with the side chains of Arg40, Arg43 and Lys103. In addition to these interactions, the crystal packing allows the formation of hydrogen-bonds between the oxygen atoms of the α-phosphate of ATP

and the side chain amino group of Lys192 from the second monomer (monomer B) of the asymmetric unit (Figure 2c). The structure of the *Aa*BPL:biotin:ATP complex reveals that the β- and γ- phosphates of ATP protrude from the adenylate binding site pointing into the inter-subunit cavity formed by the N- and C-terminal domains and are solvent-exposed (Figure 2a). In the crystal, the β-phosphate is stabilised by the hydrophilic environment formed by Asp96 from the β_5-strand and the side chains of Arg40, Lys92 and Lys103. The γ-phosphate interacts weakly with the hydroxyl group of Tyr98 and the side chain of Lys92 from the β_4- and β_5- strands respectively. Interestingly, the structure of the biotin-ATP complex indicates that the backbone carbonyl of Leu230 and the hydroxyl group of Ser229 from the β_{12}-strand of the C-terminal domain also form hydrogen-bonds with the oxygen atoms of the γ-phosphate (Figure 2c). The terminal diphosphate, is thus in a position which allows binding of Mg^{2+} resulting in the easy elimination of the pyrophosphate leaving group and the formation of biotinyl-5'-AMP. It is perhaps noteworthy that in this structure the side chain of Arg231 points towards the phosphate binding cavity. In BirA the equivalent residue (Arg317) has been shown by mutagenesis to be directly involved in MgATP binding.[28]

3. Capture of the biotin-ATP intermediate complex of AaBPL

In the *Aa*BPL and *Ph*BPL complexes with biotin and ATP, the biotin binding loops display significant conformational differences which can be attributed to the sequence differences within their respective glycine rich motifs - *Aa*BPL $_{37}$GRGRLG$_{42}$ and *Ph*BPL $_{45}$GHGRLN$_{50}$ (Figure 3). In comparison to the *Aa*BPL:biotin:ATP structure, in the *Ph*BPL complex, the loop is displaced and the side chain of the Asn50 is solventexposed and interacts weakly with the His46 (Bagautdinov and Kunishima, PDB entry: 2DTO). The Lys111 was proposed to play a key role in the formation of biotinyl-5'-AMP in *Ph*BPL by stabilising the pentacoordinate transition state of the α-phosphate before formation of the phosphoanhydride bond.[18] In the *Aa*BPL:biotin:ATP complex the equivalent conserved lysine (Lys103) adopts a similar conformation but forms long-range hydrogen-bonds with the carboxyl group of the biotin.

Of particular note are the specific orientation and the interactions made by the Arg40 residue in the structure of the *Aa*BPL complex which differ significantly from those in the *Ph*BPL and BirA substrate complexes.[14,16,18]. Arg40 plays an essential role in stabilising the biotin and ATP ligands within the active site (Figures 2b and 3). In the AaBPL:biotin:ATP ternary complex, the side chain of Arg40 interacts with the carboxyl group of the biotin and forms direct hydrogen-bonds with the oxygens of the α- and β- phosphates of ATP. The backbone NH of the Arg40 residue also interacts with the carboxyl group of the biotin. This orientation of the conserved arginine side chain has not been observed in other ligand-bound BPL structures. In the structure of the BirA:biotin complex, the backbone NH of the equivalent arginine (Arg118) forms hydrogen-bonds with the carboxyl group of the biotin and in the complex with biotinol-5'-AMP, the backbone NH interacts with the oxygen atoms of the phosphate group of the adenylate.[14,16] Thus, in the BirA complexes,

the side chain of the Arg118 appears to play mainly a structural role by closing the active site and sequestering the ligands *via* weak hydrophilic interactions. In the *Ph*BPL:biotin:ATP complex the Arg48 strongly stabilises the α-phosphate of ATP by hydrogen-bonding via its backbone NH and weakly through its sidechain.. While in the structure of the *Aa*BPL:biotin:ATP complex the side chain of the Arg40 adopts a bent conformation in order to interact with the biotin *and* the α- phosphate, in the *Ph*BPL:biotin:ATP complex, the side chain of the Arg48 points towards the γ-phosphate and has a similar orientation with respect to the triphosphate chain (Figure 3). This conformation of the Arg48 allows the side chain to form a network of weak hydrogen-bonds with the oxygen atoms of the α- and β-phosphate groups of ATP (as well as with the phosphate of the biotin in the *Ph*BPL:biotinyl-5'-AMP complex). In contrast to *Aa*BPL, neither the backbone NH nor the side chain of the Arg48 interacts with the carboxyl group of the biotin in *Ph*BPL. These combined analyses suggest that their are subtle differences in the way each BPL isozyme achieves substrate binding and

catalysis.

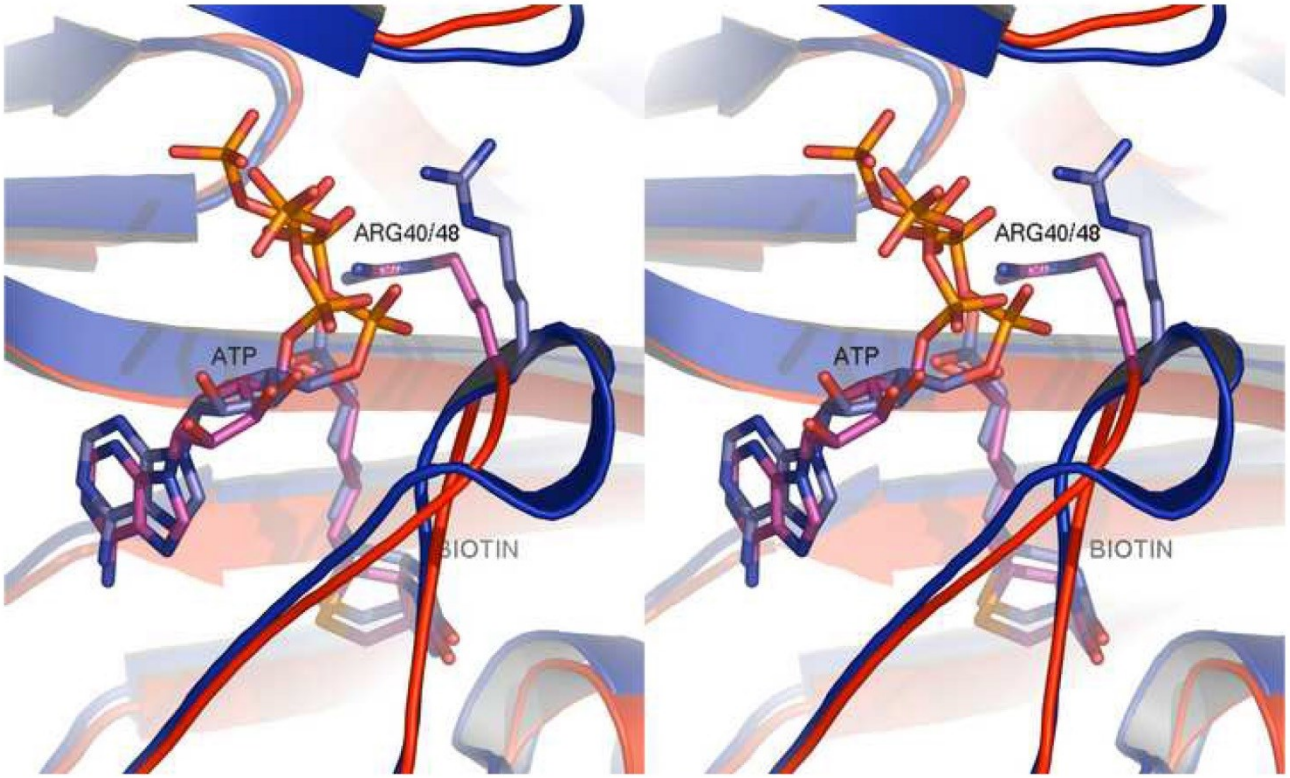


Figure 3. Comparison of the ligand-binding sites of the ATP and biotin-bound complexes of *Aa*BPL and *Ph*BPL. The *Ph*BPL PDB file used was 2DTO. The colouring is as follows: *Aa*BPL, red cartoon; ATP, biotin and Arg40 shown as sticks with purple carbons; *Ph*BPL, blue cartoon; ATP, biotin and Arg48 are shown as sticks with light blue carbons. The equivalent arginine residues and ATP are shown to be in significantly different positions.

4. The Crystal structure of the AaBPL R40G:biotin complex

Cronan and colleagues had previously noted that an *E. coli* BirA R118G mutant, previously generated by early mutational studies and analysed Kwon and Beckett, was a "promiscuous" biotinylating enzyme.[22,25,29,30] This BirA R118G is defective in binding of both biotin and biotinyl-5'-AMP whereas ATP is bound normally. The dissociation constant of the R118G mutant for biotinyl-5'-AMP binding is 400-fold greater than that of wild type BirA whereas the biotin binding constant is 100-fold reduced. In the *E. coli* enzyme this loop is important for biotin and adenylate intermediate binding only. To explore the effect of the mutation observed with BirA R118G, the conserved Arg40 residue in *Aa*BPL was similarly mutated to glycine generating an *Aa*BPL R40G mutant. This enzyme crystallised under similar conditions to the native ligase but the mutation had an impact on the growth of the crystals and significantly reduced the rapid nucleation observed with native *Aa*BPL. The crystal structure of the mutant *Aa*BPL R40G was determined in complex with biotin and the density for the R40G containing loop and for bound biotin is clear, indicating that biotin binding stabilises the loop (Figure 4a). The main conformational differences observed between the structures of the mutant and the wild-type enzymes are within the active site. The biotin binding mode of the *Aa*BPL R40G:biotin complex closely resembles that of the ATP-biotin complex of the wild type enzyme. While the *ureido* group and thiophene ring show near identical interactions with the wall formed by the ・6- and ・7- strands in both structures, in the mutant the carboxylate group of biotin is shifted towards the side chains of Lys103 and Asp96 on the β_6- and β_7- strands.

The biotin:ATP binding loop (residues 37-47) containing the R40G mutation adopts a dramatically different conformation to the one observed in the *Aa*BPL:biotin:ATP wildtype complex structure (Figure 4b). As a result, the interactions with the glycine-rich motif $_{37}$GRGGLG$_{42}$ have changed significantly and the ATP binding site is now blocked by the mutated loop. The carboxylate group of the biotin interacts *via* hydrogen-bonds with the backbone NH of the mutated Gly40, which occupies a position intermediate between the previously described ATP and biotin positions. The R40G mutation has caused the side chain of Arg38 to adopt a different orientation and point towards the oxygen atom of the biotin *ureido* group. The backbone and side chain of Leu41 have shifted towards the active site and now blocks ribose sugar binding. The strong hydrogenbond formed between the backbone NH of the Gly42 and the backbone carbonyl of the Gly40 stabilises the turn of the loop formed by the aliphatic, mutated GRGGLG glycinerich motif. While in the wild-type complex structure the side chain of the Arg43 interacts with the α-phosphate of ATP, in the structure of the mutant complex the side chain of this residue points towards the solvent. The mutated loops undergoes a single helical turn between residues Lys44 and Gln48 with the resultant repositioning of the Trp45 sidechain preventing binding of the adenine ring of ATP. Furthermore, the side chain of Leu46, which contributes to the stabilisation of ATP in the native *Aa*BPL:biotin:ATP complex, is repositioned to weakly interact with the aliphatic chain of the biotin in the structure of the mutant complex.

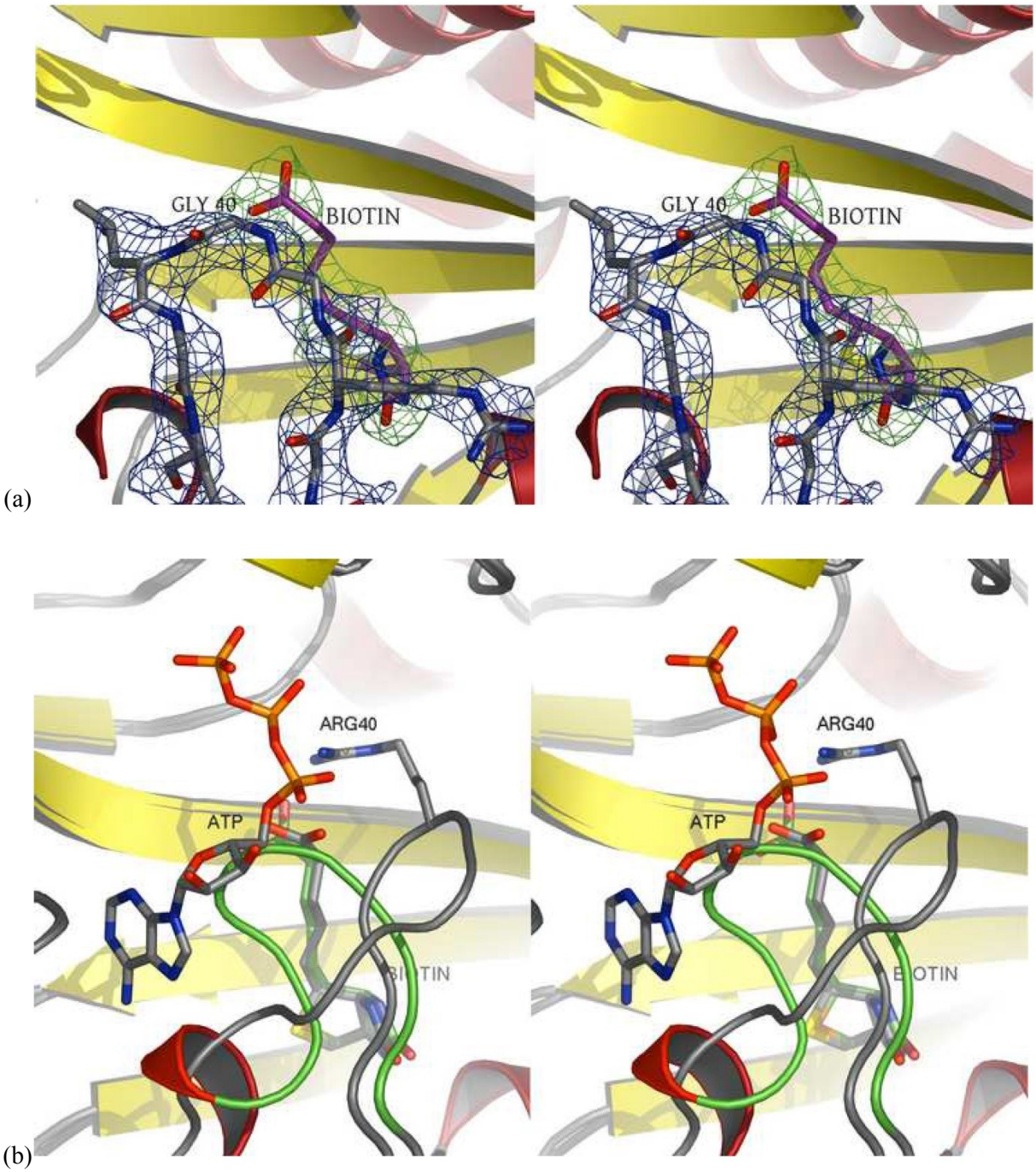


Figure 4. A close-up stereo view of the mutant R40G *Aa*BPL:biotin complex binding site. (a) Electron density is shown around the R40G containing loop (blue chickenwire) and the biotin ligand (green chickenwire). The final omit $|F_o| - |F_c|$ map contoured at 3 σ is shown. (b) Comparison of the loop position between *Aa*BPL R40G:biotin complex and wild type *Aa*BPL:biotin:ATP complex. The mutant loop and bound biotin are coloured green. ATP, biotin and Arg40 are shown as sticks in the native structure with grey coloured carbons. The mutant loop position can be clearly seen to preclude ATP binding.

5. Binding studies

Isothermal titration calorimetry (ITC) experiments were carried out to determine the thermodynamic parameters for binding of biotin and ATP with *Aa*BPL, and to investigate if the binding events occur *via* the ordered sequential mechanism established for BirA previously. Since the final step of the mechanism, the formation of biotinyl-5'- AMP, is Mg^{2+} dependent, the binding of the ligands was studied in the presence of 4 mM EDTA.[31] This is a valid simplification since the crystal structure of *Aa*BPL complexed with biotin and ATP indicates that Mg^{2+} ions are not required for binding of the substrates. The ITC results indicate that *Aa*BPL binds both substrates in an exothermic process (Figure 5a and 5b). The thermodynamic parameters (Table 1) suggest relatively weak protein-ligand interactions for both biotin and ATP binding under these conditions (K_D = 3.5 μM and KD = 7.2 μM respectively), the interaction with biotin being 1000-fold weaker than that reported for biotin with BirA (K_D = 45 nM).[22] ATP and biotin binding are enthalpy-driven with a favourable entropy change. For an ordered mechanism to occur, binding of one ligand would be prerequisite for binding of the other. However, in contrast to BirA, biotin is not required for ATP binding to *Aa*BPL. This result is unlikely to be an artefact since it could be shown by streptavidin titration that less than 4% biotin was present in the *Aa*BPL preparation and thus the amount of biotin initially present in the active site could be ignored (see supplemental data). The small difference (a factor of 2) between the K_D values for binding of ATP and biotin suggests that ligand binding to *Aa*BPL occurs *via* a random process.

Table 1. **Thermodynamic parameters derived from the titrations of *Aa*BPL and *Aa*BPL R40G with biotin and ATP before and after formation of the binary complexes formed with the two ligands.**

	K_D μM	ΔH kcal mol^{-1}	$-T\Delta S^o$ kcal mol^{-1}	ΔG^o kcal mol^{-1}
(i) *Aa*BPL + biotin	3.5 ± 0.2	-6.4 ± 0.2	-1.1 ± 0.3	-7.5 ± 0.1
(ii) *Aa*BPL + ATP	7.2 ± 1.3	-3.6 ± 0.8	-3.4 ± 0.9	-7.0 ± 0.8
(iii) *Aa*BPL:biotin + ATP	4.6 ± 0.5	-6.0 ± 0.5	-1.3 ±0 .6	-7.3 ± 0.5
(iv) *Aa*BPL:ATP + biotin	0.7 ± 0.1	-10.4 ± 0.2	2.0 ± 0.2	-8.4 ± 0.2
*Aa*BPL R40G + biotin	8.3 ± 2.1	-5.6 ± 1.5	-1.3 ± 3.8	-6.9 ± 2.3
*Aa*BPL R40G + ATP	>200	*Nd*	*Nd*	*Nd*
*Aa*BPLR40G:biotin + ATP	>200	*Nd*	*Nd*	*Nd*
*Aa*BPL R40G:ATP + biotin	15.4 ± 3.7	-5.9 ± 2.4	-0.6 ± 4.5	-6.5 ± 2.1
BirA + biotin	*(4.5 ± 0.2) x 10^{-2}*	*-9.4 ± 0.3*	*-0.5 ± 0.3*	*-9.9 ± 0.1*
BirA R118G + biotin	*1.8 ± 0.4*	*-9.3 ± 0.4*	*0.4 ± 0.4*	*-8.9 ± 0.1*

The error in ΔH is ± 5% and is mainly due to the differences in enzyme and ligand concentrations. ΔG^o is calculated from the binding determined by ITC: $\Delta G° = -RT \ln K_A = \Delta H - T.\ \Delta S^o$. $K_A = 1/K_D$**.** The parameters of BirA and BirA R118G binding to biotin are included in the table for comparison.[22, 23]

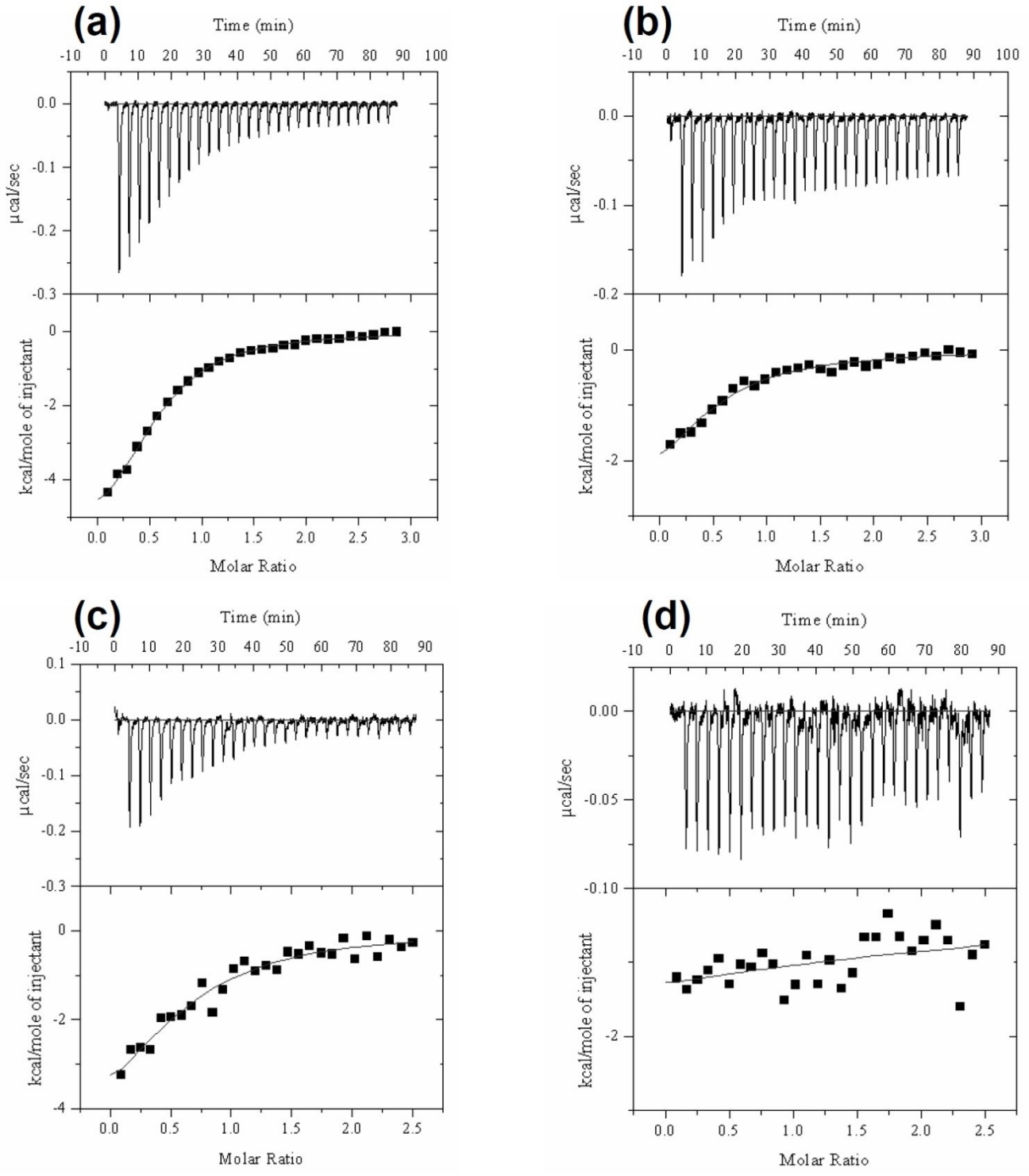


Figure 5. Isothermal titration calorimetry measurements of *Aa*BPL with (a) biotin and (b) ATP and *Aa*BPL R40G with (c) biotin and (d) ATP. The upper panels show the raw data of heat changes upon addition of ligand. The lower panels show the processed data corresponding to the heat of each injection plotted against the molar ratio of ligand to enzyme after subtraction of the heat of ligand dilution. All buffers contained 10 mM Hepes pH 7.5, 4 mM EDTA. The derived thermodynamic parameters are given in Table 1.

The cooperativity of the binding events after formation of the binary complexes *Aa*BPL:biotin and *Aa*BPL:ATP was further investigated by titration with ATP and biotin, respectively. The ITC results revealed that binding of biotin and ATP was greatly enhanced by the presence of the other substrate. The affinity (K_A) of biotin with the *Aa*BPL:ATP binary complex is five times greater than for *Aa*BPL alone (K_D = 0.7 μM compared with 3.5 μM), while the K_A of ATP for *Aa*BPL increases nearly two-fold for the *Aa*BPL:biotin complex (4.6 μM compared with 7.2 μM, Table 1). The enthalpy and entropy changes measured for the two possible pathways−i and iii, and ii and iv−leading to the formation of the *Aa*BPL:biotin:ATP ternary complex correspond within experimental error and indicate cooperativity between biotin and ATP binding (Figure 6). Despite relatively large variations in apparent enthalpy and entropy changes for the different *Aa*BPL binary and ternary complexes, the overall changes in binding free energies are much smaller and this may be attributed to the ubiquitous enthalpy‑entropy compensation frequently observed in biomolecular systems.[33]

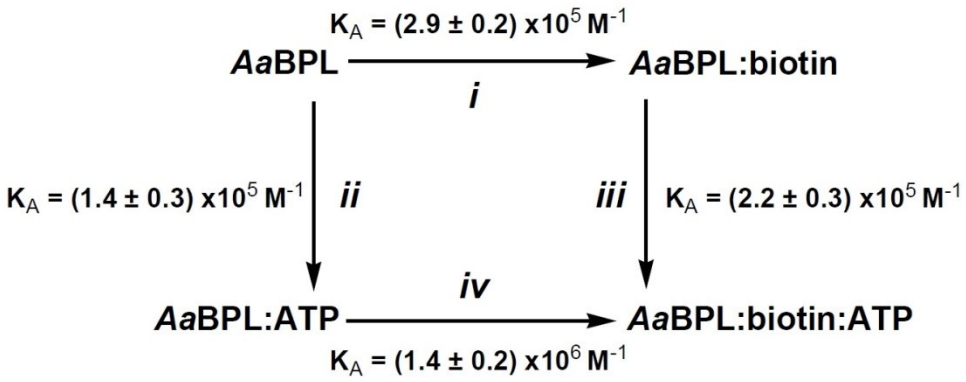


Figure 6. Formation of the ternary *Aa*BPL:biotin:ATP complex *via* the pathways i and iii, and pathways ii and iv. In *Aa*BPL, both pathways occur. K_A is the association constant ($1/K_D$).

To investigate the role of Arg40 in the ligand binding process, ITC experiments were also performed with the*Aa*BPL R40G enzyme. Titrations of this mutant under the experimental conditions used for the wild type indicated that the affinity for biotin is only 2.5-fold lower than that observed for the native enzyme (K_D 8.3 μ M compared with 3.5 μ M; see Figure 5c and Table 1). This suggests that, in contrast to the *E. coli* mutant BirA R118G, which displayed greatly reduced biotin binding, the Arg to Gly mutation has little effect on biotin binding in *A*aBPL. Thus, while the affinity of BirA for biotin is approximately 100-fold greater than that of*Aa*BPL, the dissociation constants for the biotin complexes of the two mutants BirA R118G and *Aa*BPL

R40G are similar (K_D 1.8 μ M and 8.3 μ M, respectively).[22] Both apparent enthalpy and entropy changes are favourable for biotin binding to the mutant (Table 1). Biotin was shown to bind to *Aa*BPL R40G withK_D = 15.4 μ M when the mutant was titrated in the presence of ATP. The small difference, a factor of <2, between the measured dissociation constants for the *A*aBPL R40G and the *Aa*BPL R40G binary complex suggests that the presence of ATP does not significantly alter biotin binding.

In contrast to the biotin data, titrations of the mutant enzyme with ATP showed only small heat signals, principally due to heat of dilution, indicating that binding of ATP to the *Aa*BPL R40G is negligible (K_D > 200 μ M) under these conditions (Figure 5d). The side chain of Arg40 thus appears to be necessary for ATP binding in the absence of Mg^{2+}. The structure of the mutant *Aa*BPL R40G:biotin complex revealed that biotin binding affects the conformation of the ATP-binding pocket and titration of this complex with ATP also showed very weak binding (K_D > 200 μ M).

6. Substrate biotinylation with AaBPL R40G

Previously, we showed that the C-terminal domain of *A. aeolicus* BCCP (BCCP Δ 67) is a substrate for biotinylation with *Aa*BPL.[21] Incubation of apo-BCCP Δ 67 with *Aa*BPL or the mutant *Aa*BPL R40G in the presence of biotin, ATP and $MgCl_2$ for 20 min at 65 ° C led to the formation of biotinylated-BCCP Δ 67. While, on the basis of streptavidin Western blots of the reactions, the R40G mutant appears to be a poorer catalyst than the wild type enzyme (Figure 7a, lanes 9 and 10), mass spectrometry of the biotinylated BCCP Δ 67 from the R40G-catalysed reaction revealed the formation of a singly biotinylated BCCP Δ 67 species with a mass increase equivalent to one biotin moiety (10,740.7 – 10,966.2 Da) (Figure 8c). The mutant substrate BCCP Δ 67 K117L, which lacks the target Lys117, is not biotinylated by *Aa*BPL or by *Aa*BPL R40G, showing that none of the remaining four lysine residues was derivatised and suggests a significant degree of target specificity (Figure 7a, lanes 12–14). However, Western blots of these reactions also showed that *Aa*BPL R40G, in contrast to the wild type enzyme, self-biotinylates (Figure 7a). This was confirmed by mass spectrometry analysis of the R40G species after incubation; mass increases in the range 200 ~ 2000 Da (data not shown) were observed, indicating that the single R40G mutant, which contains lysine residues, undergoes multiple biotinylation.

The equivalent BirA mutant, BirA R118G, has been shown to biotinylate itself *in vivo* and *in vitro*, and to biotinylate non-cognate acceptors such as bovine serum albumin (BSA) or RNAse A.[25] On the basis of the fact that BirA R118G binds biotinyl-5′ -AMP less tightly than wild-type BirA (K_D 20 nM and K_D 45 pM, respectively), it has been suggested that this could be a direct consequence of non-specific solution reactions of free biotinyl-5′ -AMP with exposed lysine residues. [22,25] Our suggestion that *A. aeolicus*BCCP Δ 67

K117L is not multiply biotinylated is supported by recent chemical biotinylation studies of *E. coli*BCCP87 with free biotinyl-5′ -AMP, which have demonstrated that only the target lysine is modified.[33] This specificity has been attributed to the intrinsic reactivity of the target lysine and the stabilisation of the BCCP:biotinyl-5′ -AMP complex through interactions of the *ureido* ring of the biotin moiety with residues of the "thumb" region present in both *E. coli* and *A. aeolicus* BCCP. [21,34]

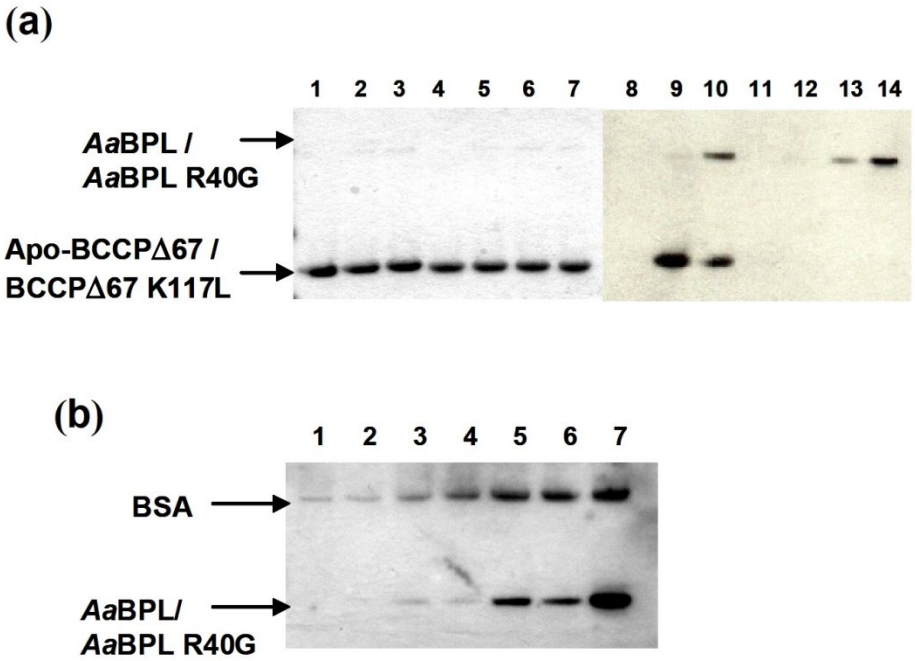


Figure 7. Streptavidin Western blot analysis of biotinylation with wild type *Aa*BPL and mutant *Aa*BPL R40G, and with BSA of (a) apo-BCCP Δ 67 and BCCP Δ 67 K117L and (b) BSA. (a) Left-hand panel: SDS-PAGE stained with Coomassie brilliant blue (lanes 1–7); right-hand panel: Western blot (lanes 8–14). Lanes 1 and 8, apo-BCCP Δ 67; lanes 2 and 9, apo-BCCP Δ 67 incubated with *Aa*BPL in the presence of biotin, ATP, $MgCl_2$ at 65 ° C for 20 min; lanes 3 and 10, apo-BCCP Δ 67 incubated with *Aa*BPL R40G under the same conditions as wild type AaBPL; lanes 4 and 11, BCCP Δ 67 K117L; lanes 5 and 12, BCCP Δ 67 K117L incubated with *Aa*BPL in the presence of biotin, ATP, $MgCl_2$ at 65 ° C; lanes 6 and 13, BCCP Δ 67 K117L incubated with*Aa*BPL R40G under the same conditions; lanes 7 and 14, BCCP Δ 67 K117L incubated with AaBPL R40G in the presence of 100 μ M biotin, ATP, $MgCl_2$ at 65 ° C. The intensity of the AaBPL R40G spot (in lane 10), given that the protein concentrations of R40G and BCCP Δ 67 are 1/10 indicates the multiple nature of R40G self-biotinylation. This was confirmed by mass spectrometry. (b) BSA (2 μ M) incubated in the presence of biotin, ATP, $MgCl_2$ at 65 ° C for 20 min with *Aa*BPL or *Aa*BPL R40G. Lane 1, BSA; lane 2, BSA with 50 nM *Aa*BPL; lane 3, BSA with 50 nM *Aa*BPL R40G; lane 4, BSA with 200 nM *Aa*BPL; lane 5, BSA with 200 nM *Aa*BPL R40G; lane 6, BSA with 500 nM *Aa*BPL; lane 7, BSA with 500 nM *Aa*BPL R40G.

To determine whether *Aa*BPL and *Aa*BPL R40G could also carry out non-specific biotinylation, the enzymes were incubated with biotin, ATP and $MgCl_2$ in the presence of (BSA) at 65 ° C. Western blot analysis of the products revealed that, while biotinylation of BSA did not occur in the presence of substrates alone or at low concentration of wild-type *Aa*BPL, the mutant *Aa*BPL R40G biotinylated itself and BSA in a manner similar to that shown for BirA R118G (Figure 7b). At high concentrations of enzyme, the wild-type *Aa*BPL also carried out self-biotinylation and non-specific biotinylation of BSA, albeit less effectively.

*Aa*BPL R40G appears to self-biotinylate preferentially rather than derivatise the substrates BSA or BCCP Δ 67 (Figure 8a and b), which is likely to be a consequence of the elevated temperature of the reaction with the thermostable enzymes. Since chemical biotinylation with biotinyl-5′ -AMP is proximity-dependent and the lifetime of AMP mixed anhydrides in aqueous solution (typically $t_{1/2} \sim 1$ min at pH 7, 25 ° C) will be significantly shorter at 65 ° C, the amount of reagent surviving to modify a non-cognate acceptor will be significantly less. [25,35]

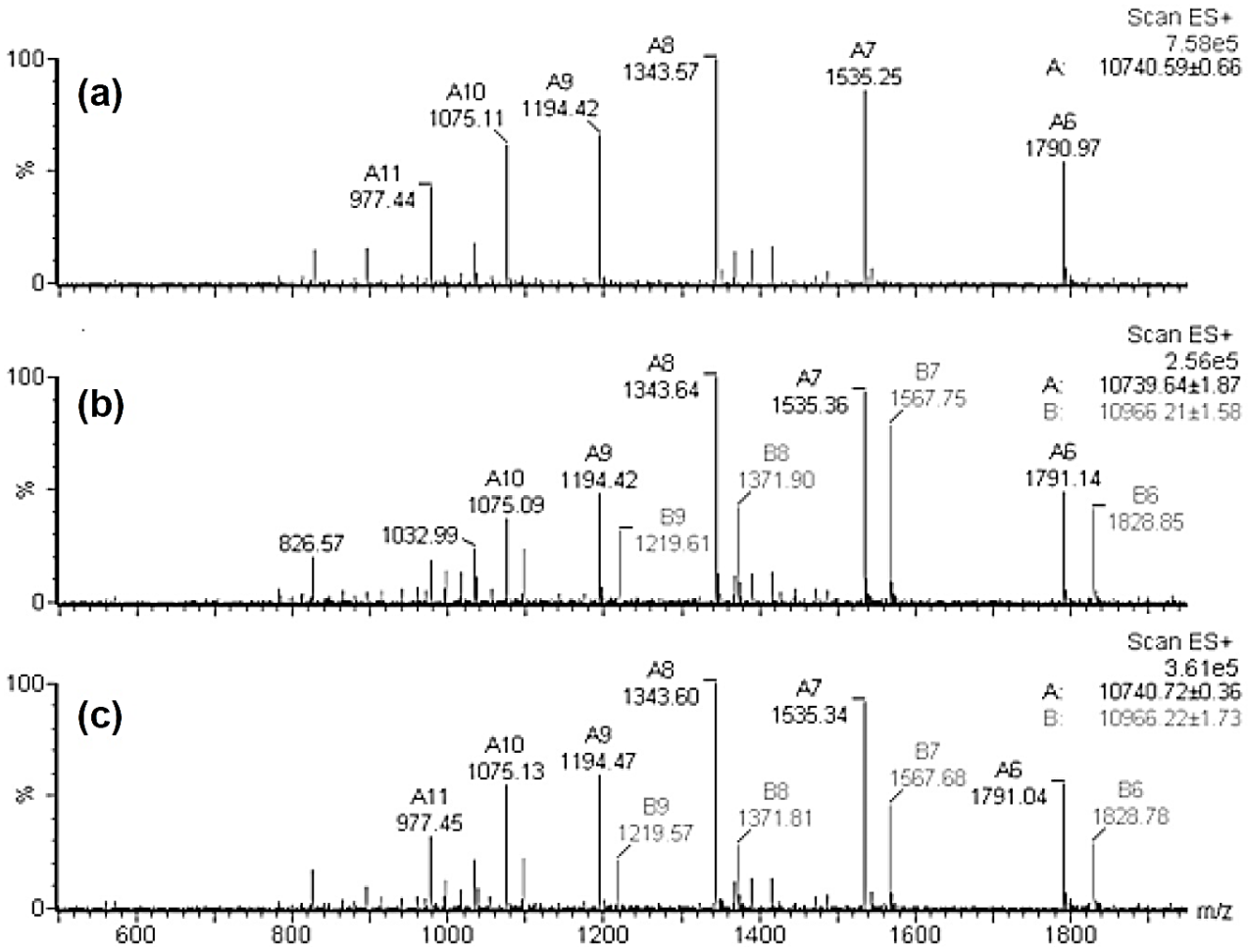


Figure 8. LC-ESI-MS analysis of apo-BCCP Δ 67: the 6+ to 10+ charge states and the deconvoluted masses. (a) apo-BCCP Δ 67 incubated in presence of biotin, ATP, $MgCl_2$ at 65 ° C for 15 min; (b) apo-BCCP Δ 67

incubated with *Aa*BPL in the same conditions; (c) apo-BCCP Δ67 incubated with *Aa*BPL R40G in the same conditions. The theoretical mass for apo-BCCP Δ67 is 10,739.63 Da and the theoretical mass for holo-BCCP Δ67 is 10,965.94 Da.

Conclusion

We have determined the crystal structures of the monomeric class I *Aa*BPL in the ligand-free form and in complex with biotin and ATP. We found that *Aa*BPL displays a high level of structural similarity with the *E. coli* BirA and *P. horikoshii* BPL enzymes with regard to the domain arrangement and overall fold. [16, 18]The geometry of the novel *Aa*BPL:biotin:ATP complex presented here suggests that it is similar to that of the transition state that precedes biotinyl-5′-AMP formation. In the structure of the *Aa*BPL:biotin:ATP complex, the biotin co-factor and the AMP moiety of ATP are buried in the adenylate-binding site, and the biotin carboxylate and the α-phosphate are located in close proximity to each other at the entrance of the active site. This suggests that subsequent nucleophilic attack of the carboxylate anion on the α-phosphate leading to the formation of the phosphoanhydride bond of the product requires only minor changes to the conformation of the active site and may be facilitated by Mg^{2+} binding to the β- and γ-phosphates of ATP, which are solvent-exposed and located in the inter-subunit domain formed by the N- and C-termini. The oxygen atoms of the ATP γ-phosphate form long-range hydrogen bonds with residues located on the SH3-like barrel C-terminal domain also characterized in the structural homologue phenylalanine tRNA synthetase.[8] This is the first demonstration that direct interactions of the substrates and the C-terminal domain of a BPL occur. This has been further highlighted in the recent co-crystal structures of *Ph*BPL R48A and K111A with BCCP Δ N76, the C-terminal domain shows the largest variations in the different stages of the complex, which suggests its functional importance in this isoform.[19]

The bent conformation of the adenylate described in the *Ph*BPL:biotinyl-5′-AMP and BirA:biotinol-5′-AMP complexes has also been observed in the structural homologue *Thermoplasma acidophilum* LplA (*Ta*LplA) bound to lipoyl-AMP. [36,37] However, the orientation of the triphosphate chain of ATP is not conserved in these enzymes. In the structure of *Ta*LplA bound to MgATP, the ATP moiety adopts a U-shaped conformation and the phosphate chain points inward towards the bifurcated lipoyl-AMP-binding pocket, which is accessible through a tunnel-like entrance. The β-phosphate occupies roughly the same position as the carboxyl group of lipoyl-AMP inside the predominantly hydrophobic cavity. The U-shaped conformation of ATP is also observed in the structures of class II aminoacyl-tRNA synthetases bound to ATP but here the β- and γ-phosphates remain solvent-exposed. For example, comparison of the structures of *Thermus thermophilus* glycyl-tRNA synthetase bound to ATP and glycyl-AMP show that the ATP phosphate chain, located at the entrance of the active site, is bent away from the binding site of the glycyl moiety. The resultant

aminoacyl-adenylate adopts an extended conformation.[38] Analysis of the structures of*Staphylococcus aureus* threonyl-tRNA synthetase bound to ATP and Zn^{2+} and to an adenylate analogue (threonyl-AMS) and the structures of the MgATP and alanine complexes of *A. aeolicus* alanyl-tRNA synthetase supports this common binding mode in class II tRNA synthetases.[39,40]

The structure of the *Aa*BPL:biotin:ATP complex has revealed the importance of Arg40 from the glycine-rich motif in the stabilisation of the two ligands. The side chain of this conserved arginine plays an important role in the neutralisation of the negative charges of the triphosphate chain of ATP as well as positioning the biotin carboxylate in close proximity to the α-phosphate. The side chain interacts with the ATP phosphate chains in two different ways in the *A. aeolicus* and *P. horoikoshii* BPL:biotin:ATP ternary complexes. In the *Aa*BPL complex, Arg40 satisfies only the negative charges of the α- and β-phosphates, and leads to a requirement for a metal ion for catalysis. Mutation of the arginine to glycine resulted in the dramatic collapse of the substrate-binding loop. The crystal structure of *Aa*BPL R40G bound to biotin shows that the mutated binding loop is ordered but blocks the ATP-binding pocket. In contrast, in the structures of the mutant PhBPL R48A complexes, the Ala48 side-chain adopts the same orientation as that of the Arg48 of the wild-type *Ph*BPL complexes and the turn of the loop containing the $_{45}$GHGRLN$_{50}$ motif shows very little conformational change in any native or mutant structures of *Ph*BPL.[19]

To investigate the effects of the R40G mutation on substrate binding in solution, comparative ITC experiments were carried out with wild-type *Aa*BPL and mutant *Aa*BPL R40G. The ITC analysis showed that*Aa*BPL binds biotin and ATP randomly, and that the binding events occur *via* a cooperative process. This contrasts with the behaviour of *E. coli* BirA, where binding of the substrates follows a sequential ordered mechanism in which formation of the BirA:biotin complex is a prerequisite for ATP binding.[16] The ordering of binding exhibited by BirA appears necessary in view of its dual role as a repressor and ligase [41,42] −it makes sense that recruitment of ATP should be the second step. In contrast, for a monofunctional BPL such as *Aa*BPL, which lacks a regulatory role, the binding order is not crucial. The BirA adenylate-binding loop, which has been shown to be involved in ATP binding and dimer packing, is disordered in the structures of the ligand-free and the biotin-bound BirA complexes. [14,15] In contrast, this loop is organised in the ligand-free form of *Aa*BPL, resulting in a well defined ATP-binding pocket. The K_m values of *Aa*BPL and BirA for ATP also reflect the greater affinity of *Aa*BPL for the nucleotide (15 μM and 200 μM, respectively) while the $K_m$values for biotin are similar (440 nM and 300 nM, respectively). [24,28] The order of binding of the substrate and ATP in the adenylate-synthesising enzymes appears to be peculiarly isozyme-specific, since the eukaryotic class IV BPL from *Saccharomyces cerevisae* binds ATP before biotin, while amino acid and ATP binding is a random process in the class II tRNA synthetases. [43,44]

The collapse of the substrate-binding loop observed in the structure of the R40G mutant adversely affects ATP binding but does not inactivate the enzyme completely (*vide infra*). We used mass spectrometry and

streptavidin Western blot studies to reveal that *Aa*BPL R40G can catalyse the formation of biotinyl-5′ -AMP in the presence of high concentrations of ATP and promiscuously biotinylate the target lysine of apo-BCCP Δ 67, and residues on BSA and the mutant itself. It can be envisaged that, as a consequence of the changed conformation of the binding site, the *Aa*BPL R40G:biotinyl-5′ -AMP binary complex can dissociate and release the adenylate in the absence of the acceptor protein. Similar behaviour has been observed with the BirA mutant R118G and in native *E. coli* methionyl-tRNA and valyl-tRNA synthetases. [25,45,46] In the structure of *Ph*BPL R48A K111A co-crystallised with biotin and ATP, the disordered triphosphate chain reduces the reactivity between the ligands and the authors subsequently used this double mutant (and the PhBPL R48A single mutant) to allow co-crystallisation with *Ph*BCCP Δ N76, the structure of which was reported during the final stages of preparation of our manuscript.[19] It is interesting to note that in the *Ph*BPL R48A K111A:*Ph*BCCP Δ N76 complex, the biotinyl domain is in the holo form and also contains bound biotin and adenosine, suggesting that the double mutant is catalytically active. Multiple conformations of the proteins were observed in the crystals obtained by Bagautdinov and colleagues, giving them a view of the enzyme throughout the catalytic process. They propose a catalytic dyad role for residues Asn103 and Asp104 (conserved as Asn95 and Asp96 in *Aa*BPL), that orientates and deprotonates the target lysine. On the other hand, the *Ph*BPL BCCP does not contain the conserved thumb found in *E. coli* and *A. aeolicus*BCCPs, and it appears that the biotinylation process in different species has subtle nuances that require detailed individual study.

Our study sheds light on the effects of mutation within the biotin-binding loop on substrate specificity of the bacterial BPLs and may provide insight into the molecular basis of human multiple carboxylase deficiency syndrome (MCD). This disease is linked to mutations in the human holocarboxylase synthetase enzyme (HCS), a class IV BPL. It has been demonstrated that the minimum functional HCS protein is retained in the 349 C-terminal amino acids, which show highly conserved sequence similarity to the putative biotin-binding domain of bacterial BPLs.[47] Further deletion of the N-terminus of the human enzyme to allow comparison with the 276 amino acids of *Aa*BPL resulted in 33.3 % sequence similarity and 19.2 % sequence identity between the human HCS domain and the *A. aeolicus* enzyme. A commonly recurring HCS mutation leading to MCD in the $_{505}$GKGRGG$_{510}$ motif is R508W which maps to Arg40 of *Aa*BPL. Patients with this missense R508W mutation are responsive to biotin and can be treated clinically with pharmaceutical doses of the vitamin. [48,49] It seems likely that further studies on the bacterial proteins may uncover details of enzyme action relevant to this area of human biochemistry.[50]

Materials and Methods

Materials

All chemicals used in the preparation of buffers were at least reagent grade. D-Biotin was purchased from Sigma. For analysis with Western blot, the blocking solution (SuperBlock® Blocking buffer in PBS) was obtained from Pierce and the antibody (Streptavidin HRP) from BD Pharmingen™. The BSA used for the *in vitro* biotinylation experiment with *Aa*BPL R40G was also obtained from Pierce. For the DNA manipulations, PCR was done using Ready To Go PCR™ beads (Amersham) and the QIAprep® Spin Miniprep kit was obtained from Qiagen. Restriction endonucleases were purchased from New England Biolabs.

Expression and purification of AaBPL

*Aa*BPL was expressed and purified essentially as described but with the following modifications.[21] The cell pellet, suspended in 10 mM Hepes pH 7.5 containing one tablet of Complete™ Proteinase Inhibitor Cocktail (Roche), was lysed by sonication (15 min of 30 s cycle). The cell lysate was clarified twice by centrifugation in the presence of DNase (Roche). The protein was then purified on a Tricorn H R 10/100 column loaded with 15S beads (Amersham) and eluted with a linear salt gradient (0–1 M NaCl in 10 mM Hepes pH 7.5).

Cloning, expression and purification of mutant AaBPL R40G

Site-directed mutagenesis was carried out with the Stratagene kit. The complementary oligonucleotides encoding the desired mutation were obtained from Sigma Genosys (*Aa*BPL R40G for: GGA AGG GGA GGA CTC GGA AGG AAG TGG CTC, *Aa*BPL R40G rev: GAG CCA CTT CCT TCC GAG TCC TCC CCT TCC).The mutant *Aa*BPL R40G was then expressed and purified as described for wild-type *Aa*BPL.

Crystallisation of apo-AaBPL, AaBPL:biotin:ATP and mutant AaBPL R40G and AaBPL R40G:biotin

The purified protein *Aa*BPL was concentrated to 5 mg/ml and crystallised by the hanging drop, vapour diffusion method at 17 ° C. Each crystallisation drop contained 2 μl of protein solution and 1 μl of the reservoir solution (0.1 M Mes, 0.2 M ammonium sulphate, 15 % (w/v) PEG 5000 mono-ethyl ether, pH 6.5). To limit nucleation due to impurities, the protein was centrifuged for 1 h at 13,000 rpm before crystallisation. The crystals appeared immediately and grew non-reproducibly in a few hours to approximately 0.5 mm in length. In the case of the *Aa*BPL:biotin:ATP complex, D-biotin (2 mM, 1 M NaOH final concentration) and ATP (5 mM, pH 7 final concentration) were added to the protein before centrifugation and crystallisation.

*Aa*BPL R40G was crystallised by the hanging drop, vapour diffusion method at 17 ° C at a concentration of 6 mg/ml.

Crystal trials of the apo form of the *Aa*BPL R40G mutant protein were done by the hanging drop, vapour diffusion method at 17 ° C. The rapid and extensive nucleation observed with wild-type *Aa*BPL was significantly reduced and it took 4–14 days for the crystals to grow. Co-crystallization of the *Aa*BPL R40G mutant (carried out at 4 ° C and 17 ° C) in the presence of 2 mM biotin and 5 mM ATP afforded a heterogeneous mixture of low-quality crystals of various shapes and lengths, while crystals of *Aa*BPL R40G co-crystallised with ATP (5 mM), indicating no electron density for the ligand. We obtained crystals of the mutant in the presence of biotin alone; each drop contained 2 μl of protein (6 mg/ml) in the presence of 2 mM D-biotin and 1 μl of the reservoir solution (0.1 M Mes, 0.2 M ammonium sulphate, 10 % (w/v) PEG 5000 mono-ethyl ether, pH 6.0) at 17 ° C. The rectangular crystals took one week to grow reproducibly and were ~0.5 mm in length.

For data collection, the AaBPL and AaBPL R40G crystals were frozen after being transferred into a cryoprotectant solution containing the reservoir solution supplemented with either 20 % (v/v) glycerol or a non-drying immersion oil (Cargille Laboratories Inc).

Data collection and structure analysis

Data were collected at station BM14 ESRF Grenoble and station 10.1 SRS Daresbury. Data were processed using the programs MOSFLM and SCALA as part of the CCP4 suite of programs and the structure was solved with the program PHASER using the structure from *P. horikoshii* (PDB code 1WPY). Initial refinement was done using the program REFMAC and finished using the program phenix.refine as part of the PHENIX package.[53,54] Manual refinement was performed using the program COOT.[55] Details of the X-ray data collection, processing and statistics are presented in Table 2.

Table 2. **X-ray data collection, processing and refinement statistics. →**

A. *Data collection and Processing.*	Apo	ATP, biotin	R40G, biotin
Space group	$P2_1$	$P2_12_12_1$	$P2_12_12_1$
Unit cell parameters *a* (Å) b (Å) c (Å) α (deg) β (deg) γ (deg)	 56.513, 61.336, 73.308, 90.0, 90.50, 90.0	 41.225 81.695 143.278 90 90 90	 41.2513 79.9277 140.1733 90 90 90
Resolution (Å) [high shell]	31.47-2.30 [2.42-2.30]	39.62-2.30 [2.42-2.30]	28.84-2.55 [2.69-2.55]
Observations [high shell]	92480 [8941]	76866 [11290]	63143 [8673]
Unique observations [high shell]	20745 [2066]	21993 [3166]	13851 [1869]
Multiplicity [high shell]	4.5 [4.3]	3.5 [3.6]	4.6 [4.6]
Completeness (%) [high shell]	92.4 [64.6]	98.9 [99.0]	88.1 [83.9]
Mean I/σ(I) [high shell]	22.4 [5.0]	10.4 [2.5]	12.6 [2.2]
R sym (%) [high shell]	3.5 [30.8]	9.7 [42.3]	9.7 [63.9]
Refinement			
Total atoms	3648	4050	3874
Solvent atoms	87	187	73
Resolution range (Å)	31.47-2.30	39.62-2.30	28.84-2.55
Rcryst (%)	21.39	23.02	23.46
Rfree (%)	28.38	26.40	28.88
r.m.s.d. from ideal			
Bond lengths (Å)	0.005	0.002	0.002
Bond angles (deg)	0.695	0.688	0.710
Ramachandran plot statistics			
Residues in most favoured regions (%)	86.8	88.0	83.8
Residues in additionally allowed regions (%)	11.1	10.8	14.7
Residues in generously allowed regions (%)	1.6	05	0.5
Residues in disallowed regions (%)	0.5	0.7	1.0

Isothermal titration calorimetry

Isothermal titration calorimetry measurements were performed on a VP-ITC calorimeter (Microcal Inc). All experiments were carried out at 25 ° C in 10 mM Hepes, 4 mM EDTA, pH 7.5. The reactants (15.4 μ M wild-type *Aa*BPL and 18 μ M mutant *Aa*BPL R40G) were placed in the 2 ml sample chamber and the substrates D-biotin and ATP (0.2 mM) in the syringe were added with 29 successive additions of 10 μ l for 20 s (with an initial injection of 1 μ l). The interval between each injection lasted 180 s. The peaks generated were corrected for ligand heat of dilution and integrated using the ORIGIN software (Microcal Inc) by plotting the values in microcalories against the molar ratio of injectant to reactant within the cell. Data were fitted using the single-site binding model. From the dissociation constant K_D and the reaction enthalpy value ΔH, the change in free Gibbs energy (ΔG^o) and entropy change (ΔS^o) can be calculated using the equation $\Delta Go = RTln(1/KD) = \Delta H - T\Delta S^o$ where R is the universal gas constant and T is the absolute temperature.

***In vitro* biotinylation studies of BSA, apo-BCCP Δ 67 and BCCP Δ 67 K117L by mutant AaBPL R40G**

All *in vitro* biotinylation experiments, unless stated otherwise, were done in 10 mM Hepes, pH 7.5 at 65 ° C for 15 or 20 min and were terminated by addition of SDS sample buffer or 0.1 % (v/v) TFA. In the biotinylation studies with BSA, 2 μ M BSA was incubated with 5 μ M d-biotin, 1 mM ATP, 2 mM $MgCl_2$ and 50, 200, or 500 nM wild-type *Aa*BPL and mutant *Aa*BPL R40G. Biotinylation experiments were also carried out with the apo form of BCCP Δ 67 and with the mutant BCCP Δ 67 K117L. The substrates apo-BCCP Δ 67 and BCCP Δ 67 K117L were expressed and purified as described previously.[21] For the streptavidin Western blot studies, the reaction mixtures contained 2 μ M substrate, 5 μ M D-biotin, 1 mM ATP, 2 mM $MgCl_2$ and 200 nM enzyme *Aa*BPL or*Aa*BPL R40G. For the mass spectrometry analysis, 10 μ M apo-BCCP Δ 67 was incubated with 100 μ M D-biotin, 1 mM ATP, 2 mM $MgCl_2$ and the reaction was initiated by addition of purified *Aa*BPL R40G to a final concentration of 2 μ M.

Streptavidin Western blot analysis

Pure proteins separated by SDS-PAGE were transferred onto a 0.2 μ m pore sized Hybond-ECL nitrocellulose membrane (Amersham) using a Trans blot Semi-Dry Transfer cell. The membrane was incubated overnight in the biotin-free SuperBlock solution at 4 ° C. Excess blocking solution was then removed by washing with phosphate buffer containing 0.1% (v/v) Tween (Sigma) four times before incubation in SuperBlock solution containing 1/50,000 Streptavidin HRP for 1 h at room temperature. The membrane was washed and dried before immunolabelled proteins were detected using the ECL enhanced

chemiluminescence detection system (Amersham) in accordance with the manufacturer's protocol. Proteins were visualised on BioMax XAR film (Kodak).

Mass spectrometry

Positive electrospray ionisation (ESI) liquid chromatography (LC-ESI-MS) was used for the characterisation of *Aa*BPL, *Aa*BPL R40G and biotinylated BCCP Δ67. Protein samples were separated using a Phenomenex C5 reverse phase column on a Waters HPLC 2690. The proteins were eluted from the column with a 5%–95% (v/v) acetonitrile gradient (containing 0.1 % TFA) and analysed on a MicroMass Platform II quadrupole mass spectrometer. The molecular mass was determined by the Transform algorithms of the Mass Lynx software (MicroMass).

Protein Data Bank accession codes

The coordinates have been submitted to the PDB. The apo form of *Aa*BPL has been assigned the PDB code3FJP, the *Aa*BPL:biotin:ATP complex has PDB code 3EFS and the *Aa*BPL R40G:biotin complex has PDB code 3EFR.

References

[1] Chapman-Smith, A. & Cronan, J. E., Jr. (1999). The enzymatic biotinylation of proteins: a post-translational modification of exceptional specificity. *Trends Biochem Sci* **24**, 359-63.

[2] Cronan, J. E., Jr. & Waldrop, G. L. (2002). Multi-subunit acetyl-CoA carboxylases. *Prog Lipid Res* **41**, 407-35.

[3] Nenortas, E. & Beckett, D. (1996). Purification and characterization of intact and truncated forms of the Escherichia coli biotin carboxyl carrier subunit of acetyl- CoA carboxylase. *J Biol Chem* **271**, 7559-67.

[4] Tong, L. (2005). Acetyl-coenzyme A carboxylase: crucial metabolic enzyme and attractive target for drug discovery. *Cell Mol Life Sci* **62**, 1784-803.

[5] Lane, M. D., Rominger, K. L., Young, D. L. & Lynen, F. (1964). The Enzymatic Synthesis of Holotranscarboxylase from Apotranscarboxylase and (+)-Biotin. Ii. Investigation of the Reaction Mechanism. *J Biol Chem* **239**, 2865-71.

[6] Fujiwara, K., Toma, S., Okamura-Ikeda, K., Motokawa, Y., Nakagawa, A. & Taniguchi, H. (2005). Crystal structure of lipoate-protein ligase A from *Escherichia coli*. Determination of the lipoic acid-binding site. *J. Biol. Chem.* **239**, 2865-2871.

[7] Artymiuk, P. J., Rice, D. W., Poirrette, A. R. & Willet, P. (1994). A tale of two synthetases. *Nat. Struct. Biol.* **1**, 758-760.

[8] Safro, M. & Mosyak, L. (1995). Structural similarities in the noncatalytic domains of phenylalanyl-tRNA and biotin synthetases. *Protein Sci.* **4**, 2429-2432.

[9] Mukhopadhyay, B., Purwantini, E., Kreder, C. L. & Wolfe, R. S. (2001). Oxaloacetate synthesis in the methanarchaeon *Methanosarcina barkeri*: pyruvate carboxylase genes and a putative *Escherichia coli*-type bifunctional biotin protein ligase gene (bpl/birA) exhibit a unique organization. *J Bacteriol* **183**, 3804-10.

[10] Beckett, D. & Matthews, B. W. (1997). *Escherichia coli* repressor of biotin biosynthesis. *Methods Enzymol.* **279**, 362-376.

[11] Eisenstein, E. & Beckett, D. (1999). Dimerization of the *Escherichia coli* biotin repressor: corepressor function in protein assembly. *Biochemistry* **38**, 13077- 13084.

[12] Beckett, D. (2007). Biotin sensing: universal influence of biotin status on transcription. *Ann. Rev. Genet.* **41**, 443-464.

[13] Brown, P. H., Cronan, J. E., Grotli, M. & Beckett, D. (2004). The biotin repressor: modulation of allostery by corepressor analogs. *J. Mol. Biol.* **337**, 857-869.

[14] Weaver, L. H., Kwon, K., Beckett, D. & Matthews, B. W. (2001). Corepressorinduced organization and assembly of the biotin repressor: a model for allosteric activation of a transcriptional regulator. *Proc. Natl Acad. Sci. USA* **98**, 6045-6050.

[15] Wilson, K. P., Shewchuk, L. M., Brennan, R. G., Otsuka, A. J. & Matthews, B. W. (1992). *Escherichia coli* biotin holoenzyme synthetase/bio repressor crystal structure delineates the biotin- and DNA-binding domains. *Proc. Natl Acad. Sci. USA* **89**, 9257-61.

[16] Wood, Z. A., Weaver, L. H., Brown, P. H., Beckett, D. & Matthews, B. W. (2006). Co-repressor induced order and biotin repressor dimerization: a case for divergent followed by convergent evolution. *J. Mol. Biol.* **357**, 509-523.

[17] Xu, Y. & Beckett, D. (1994). Kinetics of biotinyl-5'-adenylate synthesis catalysed by the *Escherichia coli* repressor of biotin biosynthesis and the stability of the enzyme-product complex. *Biochemistry* **33**, 7354-7360.

[18] Bagautdinov, B., Kuroishi, C., Sugahara, M. & Kunishima, N. (2005). Crystal structures of biotin protein ligase from *Pyrococcus horikoshii* OT3 and its complexes: structural basis of biotin activation. *J. Mol. Biol.* **353**, 322-333.

[19] Bagautdinov, B., Matsuura, Y., Bagautdinova, S. & Kunishima, N. (2008). Protein biotinylation visualized by a complex structure of biotin protein ligase with a substrate. *J. Biol. Chem.* **283**, 14739-14750.

[20] Weaver, L. H., Kwon, K., Beckett, D. & Matthews, B. W. (2001). Competing protein:protein interactions are proposed to control the biological switch of the *E. coli* biotin repressor. *Protein Sci.* **10**, 2618-2622.

[21] Clarke, D. J., Coulson, J., Baillie, R. & Campopiano, D. J. (2003). Biotinylation in the hyperthermophile *Aquifex aeolicus*. *Eur. J. Biochem.* **270**, 1277-1287.

[22] Kwon, K. & Beckett, D. (2000). Function of a conserved sequence motif in biotin holoenzyme synthetases. *Protein Sci.* **9**, 1530-1539.

[23] Kwon, K., Streaker, E. D. & Beckett, D. (2002). Binding specificity and the ligand dissociation process in the *E. coli* biotin holoenzyme synthetase. *Protein Sci.* **11**, 558-570.

[24] Kwon, K., Streaker, E. D., Ruparelia, S. & Beckett, D. (2000). Multiple disordered loops function in corepressor-induced dimerization of the biotin repressor. *J. Mol. Biol.* **304**, 821-833.

[25] Choi-Rhee, E., Schulman, H. & Cronan, J. E. (2004). Promiscuous protein biotinylation by *Escherichia coli* biotin protein ligase. *Protein Sci.* **13**, 3043-3050.

[26] Dupuis, L., Leon-Del-Rio, A., Leclerc, D., Campeau, E., Sweetman, L., Saudubray, J. M.*, et al.* (1996). Clustering of mutations in the biotin-binding region of holocarboxylase synthetase in biotin-responsive multiple carboxylase deficiency. *Hum. Mol. Genet.* **5**, 1011-1016.

[27] Suzuki, Y., Yang, X., Aoki, Y., Kure, S. & Matsubara, Y. (2005). Mutations in the holocarboxylase synthetase gene HLCS. *Hum. Mutat.* **26**, 285-290.

[28] Chapman-Smith, A., Mulhern, T. D., Whelan, F., Cronan, J. E. J. & Wallace, J. C. (2001). The C-terminal domain of biotin protein ligase from *E. coli* is required for catalytic activity. *Protein Sci.* **10**, 2608-2617.

[29] Barker, D. F. & Campbell, A. M. (1981). Genetic and biochemical characterization of the birA gene and its product: evidence for a direct role of biotin holoenzyme synthetase in repression of the biotin operon in *Escherichia coli*. *J. Mol. Biol.* **146**, 469-492.

[30] Buoncristiani, M. R., Howard, P. K. & Otsuka, A. J. (1986). DNA-binding and enzymatic domains of the bifunctional biotin operon repressor (BirA) of *Escherichia coli*. *Gene* **44**, 255-261.

[31] Perozzo, R., Jelesarov, I., Bosshard, H. R., Folkers, G. & Scapozza, L. (2000). Compulsory order of substrate binding to herpes simplex virus type 1 thymidine kinase. A calorimetric study. *J. Biol. Chem.* **275**, 16139-16145.

[32] Cooper, A., Johnson, C. M., Lakey, J. H. & Nollmann, M. (2001). Heat does not come in different colours: entropy-enthalpy compensation, free energy windows, quantum confinement, pressure perturbation calorimetry, solvation and the multiple causes of heat capacity effects in biomolecular interactions. *Biophys. Chem.* **93**, 215-230.

[33] Streaker, E. D. & Beckett, D. (2006). Nonenzymatic biotinylation of a biotin carboxyl carrier protein: unusual reactivity of the physiological target lysine. *Protein Sci.* **15**, 1928-1935.

[34] Solbiati, J., Chapman-Smith, A. & Cronan, J. E., Jr. (2002). Stabilization of the biotinoyl domain of *Escherichia coli* acetyl-CoA carboxylase by interactions between the attached biotin and the protruding "thumb" structure. *J. Biol. Chem.* **277**, 21604-21609.

[35] Pavela-Vrancic, M., Dieckmann, R., Dohren, H. V. & Kleinkauf, H. (1999). Editing of non-cognate aminoacyl adenylates by peptide synthetases. *Biochem. J.* **342**, 715-719.

[36] Kim, D. J., Kim, K. H., Lee, H. H., Lee, S. J., Ha, J. Y., Yoon, H. J. & Suh, S. W. (2005). Crystal structure of lipoate-protein ligase A bound with the activated intermediate: insights into interaction with lipoyl domains. *J. Biol. Chem.* **280**, 38081-38089.

[37] McManus, E., Luisi, B. F. & Perham, R. N. (2006). Structure of a putative lipoate protein ligase from *Thermoplasma acidophilum* and the mechanism of target selection for post-translational modification. . *J. Mol. Biol.* **356**, 625-637.

[38] Arnez, J. G., Dock-Bregeon, A. C. & Moras, D. (1999). Glycyl-tRNA synthetase uses a negatively charged pit for specific recognition and activation of glycine. *J. Mol. Biol.* **286**, 1449-1459.

[39] Swairjo, M. A. & Schimmel, P. R. (2005). Breaking sieve for steric exclusion of a noncognate amino acid from active site of a tRNA synthetase. *Proc. Natl Acad. Sci. USA* **102**, 988-999.

[40] Torres-Larios, A., Sankaranarayanan, R., Rees, B., Dock-Bregeon, A. C. & Moras, D. (2003). Conformational movements and cooperativity upon amino acid, ATP and tRNA binding in threonyl-tRNA synthetase. *J. Mol. Biol.* **331**, 201-211.

[41] Streaker, E. D. & Beckett, D. (2006). The biotin regulatory system: kinetic control of a transcriptional switch. *Biochemistry* **45**, 6417-6425.

[42] Streaker, E. D., Gupta, A. & Beckett, D. (2002). The biotin repressor: thermodynamic coupling of corepressor binding, protein assembly, and sequencespecific DNA binding. *Biochemistry* **41**, 14263-14271.

[43] Pan, F., Lo, K. Y., Pai, S. H. & Lee, H. H. (1982). Kinetic mechanism of threonyltRNA synthetase from human placenta. *Int. J. Pept. Protein Res.* **20**, 159-166.

[44] Polyak, S. W., Chapman-Smith, A., Brautigan, P. J. & Wallace, J. C. (1999). Biotin protein ligase from Saccharomyces cerevisiae. The N-terminal domain is required for complete activity. *J Biol Chem* **274**, 32847-54.

[45] Gillet, S., Hountondji, C., Schmitter, J. M. & Blanquet, S. (1997). Covalent methionylation of *Escherichia coli* methionyl-tRNA synthethase: identification of the labeled amino acid residues by matrix-assisted laser desorption-ionization mass spectrometry. *Protein Sci* **6**, 2426-35.

[46] Hountondji, C., Lazennec, C., Beauvallet, C., Dessen, P., Pernollet, J. C., Plateau, P. & Blanquet, S. (2002). Crucial role of conserved lysine 277 in the fidelity of tRNA aminoacylation by *Escherichia coli* valyl-tRNA synthetase. *Biochemistry* **41**, 14856-14865.

[47] Campeau, E. & Gravel, R. A. (2001). Expression in *Escherichia coli* of N- and Cterminally deleted human holocarboxylase synthetase. Influence of the N-terminus on biotinylation and identification of a minimum functional protein. *J. Biol. Chem.* **276**, 12310-12316.

[48] Dupuis, L., Campeau, E., Leclerc, D. & Gravel, R. A. (1999). Mechanism of biotin responsiveness in biotin-responsive multiple carboxylase deficiency. *Mol. Genet. Metab.* **66**, 80-90.

[49] Morrone, A., Malvagia, S., Donati, M. A., Funghini, S., Ciani, F., Pela, I.*, et al.* (2002). Clinical findings and biochemical and molecular analysis of four patients with holocarboxylase synthetase deficiency. *Am. J. Med. Genet.* **111**, 10-18.

[50] Pendini, N. R., Bailey, L. M., Booker, G. W., Wilce, M. C., Wallace, J. C. & Polyak, S. W. (2008). Microbial biotin protein ligases aid in understanding holocarboxylase synthetase deficiency. *Biochim. Biophys. Acta.* **1784**, 973-982.

[51] CCP4. (1994). The CCP4 suite: Programs for Protein Crystallography. *Acta Crystallographica Section D* **50**, 760-763.

[52] McCoy, A. J., Grosse-Kunstleve, R. W., Storoni, L. C. & Read, R. J. (2004). Likelihood-enhanced fast translation functions. *Acta Crystallog. sect. D* **61**, 458-464.

[53] Adams, P. D., Grosse-Kunstleve, R. W., Hung, L. W., Ioerger, T. R., McCoy, A. J., Moriarty, N. W.*, et al.* (2002). PHENIX: building new software for automated crystallographic structure determination. *Acta Cryst. D Biol. Crystallogr.* **58**, 1948-1954.

[54] Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallog. sect. D* **53**, 240-255.

[55] Emsley, P. & Cowtan, K. (2004). Coot: model-building tools for molecular graphics. *Acta Crystallog. sect. D* **60**, 2126-2132.